



**UNITED STATES DEPARTMENT OF COMMERCE**  
**Bureau of the Census**  
Washington, DC 20233-0001

December 7, 2000

DSSD CENSUS 2000 PROCEDURES AND OPERATIONS MEMORANDUM SERIES R-39

MEMORANDUM FOR      Maureen Lynch  
Assistant Division Chief, Coverage Measurement Processing  
Decennial Statistical Studies Division

From:                    Donna Kostanich *OK*  
Assistant Division Chief, Sampling and Estimation  
Decennial Statistical Studies Division

Prepared by:            *Jef* Ryan Cromar, *Far* James Farber, and *Rf* Roxanne Feldpausch  
Decennial Statistical Studies Division

Subject:                Accuracy and Coverage Evaluation Survey: Final Specification for  
E-Sample Identification

I.      INTRODUCTION

This specification describes the identification of E-sample housing units (HUs) for the Accuracy and Coverage Evaluation (A.C.E.) survey. The identification of the E sample is required for person matching and final HU matching. In general, E-sample HUs are HUs enumerated in the census in block clusters or segments of clusters that were in the P sample. This specification covers only the identification of E-sample HUs. The E-sample people are all data-defined people enumerated in the census in E-sample HUs, and are identified when the person matching files are created.

The E sample is used primarily for A.C.E. estimation. The E sample and P sample will be matched to estimate how many people and HUs were omitted from or erroneously included in the census. The P sample was created through the A.C.E. sampling process, including listing sample selection, the A.C.E. reduction, small block cluster subsampling, and within-large block cluster subsampling, and consists of all HUs and people interviewed in the independent A.C.E. sample. The P-sample HUs are contained on the subsampled preliminary enhanced list (SPEL), an output of large block cluster subsampling.

The A.C.E. HUs were originally identified through the independent listing operation and placed on the independent list. Census HUs were identified through a number of census operations and were placed on the Decennial Master Address File (DMAF). During

initial HU matching and follow-up, the two address lists were compared and the results recorded to the preliminary enhanced list. The preliminary enhanced list contains the A.C.E. and census HUs that match to each other along with HUs from each list that do not match but were found to be valid HUs during field follow-up. The SPEL is the preliminary enhanced list updated with the results of large block cluster subsampling. The same HUs are on both the preliminary enhanced list and the SPEL. The P sample consists of A.C.E. HUs on the SPEL in block clusters or segments of clusters selected for interview. Census HUs may also be on the SPEL, but they are ineligible for interview.

The E sample is identified to maximize the overlap between the E sample and P sample. The overlapping is not required for A.C.E. estimation, but it increases efficiency by reducing field follow-up workloads. Field follow-up is also reduced by the E-sample subsampling described in this specification. When the number of E-sample eligible HUs in a block cluster is too large, a subsample of these HUs will be selected for inclusion in the E sample.

This specification is ordered into the following sections:

- Definitions
- Assumptions
- Overview
- Input Files
- Output Files
- Process
- Verification
- References

We state that E-sample identification is complete. This specification reflects the process as it was actually implemented. The original specification and the results of E-sample identification are documented in references 1 and 2, respectively. This final specification includes changes to the original that were required due to issues that arose during production. In general, this final version is nearly identical to the original specification.

Any questions or comments regarding this specification should be directed to Frances Farber (301-457-4282) or Deborah Fenstermaker (301-457-4195) of the Decennial Statistical Studies Division (DSSD).

## II. DEFINITIONS

### A. A.C.E. Housing Unit

A housing unit on the SPEL with an after follow-up match code (see below) of M, MU, CI, or UI. In general, these are HUs found during A.C.E. independent listing.

### B. After Follow-up Match Codes

Codes assigned to HUs during initial A.C.E. HU matching and follow-up. For the purposes of this specification, the only match codes that need to be defined are those that occur on the SPEL. As documented in reference 3, these match codes are:

M	=	The A.C.E. and census addresses match.
MU	=	The A.C.E. and census addresses match and there is not enough information on the follow-up form to confirm this match as an HU with certainty. The follow-up interview was not done, was incomplete, was never sent, had contradictory information, or was a noninterview.
UI	=	Not enough information on the follow-up form to assign a code to the nonmatched A.C.E. HU with certainty. The follow-up interview was not done, was incomplete, was never sent, had contradictory information, or was a noninterview.
UE	=	Not enough information on the follow-up form to assign a code to the census nonmatched HU with certainty. The follow-up interview was not done, was incomplete, was never sent, had contradictory information, or was a noninterview.
CI	=	The A.C.E. HU existed as an HU at the time of the follow-up interview and is correctly geocoded in the block cluster. The HU is not found in the census.
CE	=	The census HU existed as an HU at the time of the follow-up interview and is correctly geocoded in the block cluster. The HU is not found in the A.C.E.

### C. American Indian Reservation Block Cluster

A block cluster at least partially on an American Indian Reservation (AIR), according to the AIR definitions used at the time of A.C.E. block clustering. The AIR block clusters have an American Indian Country Indicator (AICIND) = 1 on the Sample Design File.

#### D. Corresponding HU

A census HU on the E-Sample Identification Input File that corresponds to an SPEL HU. Census HUs correspond to SPEL HUs through the Census Identification Number (CID), a code assigned to all census HUs. During initial HU matching and follow-up, SPEL HUs that matched to DMAF HUs were assigned the CID of the matching DMAF HU. If the DMAF HU was not deleted in a later operation, such as nonresponse follow-up, and made it onto the Hundred Percent Census Unedited File (HCUF), the source of the E sample, then the CID link still exists between the SPEL HU and the HCUF HU. The SPEL HU has the same CID as the corresponding HCUF HU. Since both A.C.E. and supplemental HUs are on the SPEL, an HCUF HU may correspond to either type of SPEL HU.

The correspondence status of an HCUF HU is denoted by the variable Match Status (MSTATUS), where

0	=	The HCUF HU is in a cluster with fewer than 80 HCUF HUs.
1	=	The HCUF HU is corresponding in a cluster with 80 or more HCUF HUs.
2	=	The HCUF HU is non-corresponding in a non-special case cluster with 80 or more HCUF HUs.
3	=	The HCUF HU is non-corresponding and in a special case cluster with 80 or more HCUF HUs.

#### E. E Sample

Census HUs in block clusters or block cluster segments that were selected for the A.C.E. interview sample. The E sample can also refer to the people enumerated in the census in E-sample HUs. The HCUF is the source of the E sample.

#### F. E-Sample Indicator

Indicates whether or not an HCUF HU is in the E sample.

1	=	The HCUF HU is in the E sample.
2	=	The HCUF HU is not in the E sample.

G. E-Sample Probability Code

Indicates which E-sample weight applies to an E-sample HU.

- |   |   |                                      |
|---|---|--------------------------------------|
| 1 | = | WEIGHTE1 applies to the E-sample HU. |
| 2 | = | WEIGHTE2 applies to the E-sample HU. |

It is possible for two weights to apply to the E sample in a single block cluster if the cluster requires E-sample subsampling.

H. HU Group

Indicates how to calculate the weight for each E-sample HU and in which set of summary HU counts each E-sample HU should be tallied. Attachment A contains a description of the HU Groups.

I. Hundred Percent Census Unedited File Housing Unit

An HU found to be a valid HU during the census and thus retained from the DMAF onto the HCUF. The HUs on the E-Sample Identification Input File are also called HCUF HUs because the E-sample input file is an extract of the HCUF.

J. Non-Corresponding Housing Unit

A census HU on the E-Sample Identification Input File whose CID was not assigned to any SPEL HUs during initial HU matching. Examples of non-corresponding HCUF HUs include census adds and previously mis-geocoded HUs that are moved into the cluster.

K. P sample

For the purposes of this specification, the P-sample is defined as the HUs in the A.C.E. interview sample. The P sample can also refer to the people interviewed in the A.C.E. in P-sample HUs. The P-sample HUs were initially identified during A.C.E. independent listing.

L. Preliminary E-Sample Indicator

Indicates whether or not an HCUF HU is initially eligible for the E sample. Some HCUF HUs may initially be eligible for the E sample but will not end up in the final E sample due to E-sample subsampling.

0	=	The HCUF HU is not initially eligible for the E sample and is not subject to E-sample subsampling.
1	=	The HCUF HU is initially eligible for the E sample and may be subject to E-sample subsampling.

M. Segment Identifier

A variable created during large block cluster subsampling and mapped onto HCUF HUs. The field assignment segments in the control number are not the segment identifiers referred to in this specification. See reference 4 for information on how segment identifiers were created in large block cluster subsampling.

N. Supplemental HU

A housing unit on the SPEL with an after follow-up match code of UE or CE. Supplemental HUs are not eligible for the A.C.E. interview sample, but were assigned segment identifiers and special interview codes during large block cluster subsampling to facilitate E-sample identification. The source of supplemental HUs was the version of the DMAF used for initial HU matching. Supplemental HUs could not be matched to any A.C.E. HUs during initial HU matching but were found to exist during HU follow-up.

III. ASSUMPTIONS

- A. Supplemental HUs were map spotted during initial HU follow-up and thus can be correctly located near their geographic neighbors on the SPEL.
- B. HCUF HUs correspond to SPEL HUs through the CID. The CIDs of corresponding HCUF HUs were assigned to their matching SPEL HUs during initial HU matching.
- C. There is no E-sample subsampling in AIR clusters, where such clusters were defined using the AIR definitions known at the time of A.C.E. block clustering.

- D. The E-sample people will be identified during file creation for A.C.E. person matching and are beyond the scope of this specification. The E-sample people are all data-defined people on the HCUF in E-sample HUs (see reference 5).
- E. All decimal numbers are rounded to six decimal places at the time of creation using standard rounding procedures, unless otherwise noted. In this specification, standard rounding means that a number with a seventh-decimal value of five or higher is rounded up in the sixth decimal. Otherwise, the sixth decimal value is unchanged.

#### IV. PROCESS OVERVIEW

The E-sample identification process described in this specification is used to identify which census HUs are in the E sample. The E-sample HUs and the people in those HUs will be matched to the HUs and people in the P sample, the A.C.E. interview sample. Those matching results form the basis of A.C.E. person and final HU estimation. This overview details the E-sample identification process. The steps in this section correspond to the steps in section VII, which contains the programming instructions and is significantly less detailed than this overview. Attachment B contains a flowchart of the E-sample identification process.

- A. The first step is to screen out clusters where the identification of the E sample is simple. The source of E-sample HUs is an extract of the HCUF called the E-Sample Identification Input File, which contains all census HUs in A.C.E. sample block clusters. The HCUF HUs not in A.C.E. clusters are excluded from the E-Sample Identification Input File, as are all group quarters records and all person records.<sup>1</sup>

If the number of HCUF HUs in a cluster is less than 80, then all of those HUs are in the E sample. Likewise, if a cluster is an AIR cluster, using the AIR definitions known at the time of A.C.E. block clustering, then all HCUF HUs are in the E sample regardless of their number. For these clusters, E-sample identification is simple. Information for these clusters is saved to the E-Sample Identification Input File and the Sample Design File, and they are finished with the E-sample identification process. Clusters that have at least 80 HCUF HUs and are not an AIR cluster continue to the next step.

- B. There are four types of clusters that are considered special cases in the E-sample identification process because all of their HCUF HUs are non-corresponding. A

---

<sup>1</sup>In this specification, HUs on the E-Sample Identification Input File are called HCUF HUs.

non-corresponding HCUF HU did not have its CID assigned to an SPEL HU during initial HU matching. For example, all census HUs added to the HCUF after initial HU matching are non-corresponding. The four types of special case clusters are:

- List/Enumerate clusters, which did not go through initial HU matching since they did not have any DMAF HUs at that point.
- Relisted clusters, which also did not go through initial HU matching due to timing constraints.
- Clusters that went through initial HU matching and had no SPEL HUs that were assigned a CID. These are clusters where all of the SPEL HUs have match codes of UI or CI.
- Clusters that had zero A.C.E. HUs designated for interview but have 80 or more HCUF HUs.

All HCUF HUs in special case clusters are initially eligible for the E sample. In special case clusters with at least 80 HCUF HUs, E-sample subsampling is required, and these clusters proceed immediately to step E. Special case clusters with fewer than 80 HCUF HUs were screened out in step A.

- C. In non-special case clusters with at least 80 HCUF HUs, the next step is to map the results of large block cluster subsampling onto the HCUF and identify those HUs that are in the E sample, out of the E sample, or have an unknown E-sample status. Corresponding HUs are HCUF HUs whose CIDs were assigned to SPEL HUs during initial HU matching. The E-sample status of corresponding HUs can be fully determined in this step using the results of large block cluster subsampling. Non-corresponding HUs have an unknown E-sample status at this point.
- If large block cluster subsampling did not occur in the cluster, then all corresponding HCUF HUs are in the E sample.
  - If large block cluster subsampling did occur in the cluster, then corresponding HCUF HUs in segments selected for the P sample are in the E sample. Corresponding HCUF HUs in segments not selected for the P sample are not in the E sample.
  - For non-corresponding HCUF HUs, step D and possibly step E will determine which are in the E sample.
- D. This step determines whether E-sample subsampling is required in a non-special case cluster. Non-corresponding HUs were not available for large block cluster subsampling and thus received no segment identifiers. To determine if E-sample subsampling is necessary, the first step is to assign segment identifiers to non-corresponding HUs. This is done by sorting all HCUF HUs geographically and



assigning non-corresponding HUs the segment identifier of the nearest previous corresponding HU.

Then, non-corresponding HUs in segments not selected for the P sample are out of the E sample. Non-corresponding HUs in P-sample segments are initially eligible for the E sample, but may be subsampled. If there are fewer than 80 E-sample eligible non-corresponding HUs, then all of them are in the E sample. If there are 80 or more, then a subsample is taken in step E.

- E. Only clusters with 80 or more E-sample eligible HUs, including special case clusters, go through E-sample subsampling. The E-sample subsampling process is a standard systematic sample with a random start and a geographic sort of HCUF HUs. The take-every is the number of E-sample eligible HUs in the cluster divided by 40, with a maximum take-every of 4. The E-sample eligible HUs that are selected in the subsample are in the E sample, while non-selected HUs are out of the E sample.
- F. The last step of E-sample identification is to update files and compute summary counts. Note that the E-sample HUs within a single cluster may have different E-sample weights. If E-sample subsampling did not occur in a cluster, then the E-sample HUs in that cluster have only one weight. However, two E-sample weights may be required in clusters that were subsampled. In non-special case clusters where E-sample subsampling occurred, corresponding HUs have a different weight than non-corresponding HUs. Only the latter HUs were subject to E-sample subsampling and thus received additional weight. In special case clusters, only one weight applies to all E-sample HUs even if E-sample subsampling occurred because all HUs in special case clusters with 80 or more HUs are subject to E-sample subsampling.

## V. INPUT FILES

### A. Subsampled Preliminary Enhanced List

Description: The SPEL is the preliminary enhanced list updated with the results of large block cluster subsampling. All HUs with the after follow-up match codes given in Section II.B are on the SPEL whether or not they were in segments selected for the P sample. There is one SPEL file for each A.C.E. regional office.

Level: Housing Unit

Scope: One record for each HU in A.C.E. sample block clusters following small block cluster subsampling.

Layout: See Attachment C.

B. E-Sample Identification Input File

Description: The E-Sample Identification Input File is an extract of the HCUF, the file that contains the data for all HUs, group quarters, and data-defined people enumerated in the census. The E-Sample Identification Input File includes certain variables from HCUF HU records but excludes person and group quarters records, and limits the geographic scope only to A.C.E. sample clusters. There is one E-Sample Identification Input File for each cluster.

Level: Housing Unit

Scope: One record for each census HU in A.C.E. sample clusters.

Layout: See Attachment G.

C. Sample Design File Version 5

Description: Version 5 of the A.C.E. Sample Design File, which reflects the previous A.C.E. sampling operations: listing sample selection, A.C.E. reduction, small block cluster subsampling, large block cluster subsampling, and targeted extended search sampling. There are 29,717 records on the Sample Design File. The name of version 5 is SDF.US5.

Level: Block Cluster

Scope: One record for each block cluster selected in the first step of the A.C.E. listing sample.

File Layout: See Attachment D.

VI. OUTPUT FILES

A. Sample Design File Version 6

Description: Version 6 of the A.C.E. Sample Design File, which includes the results of E-sample identification. Version 6 will be named SDF.US6.

Level: Block Cluster

Scope: One record for each block cluster selected in the first step of the A.C.E. listing sample.

File Layout: See Attachment D.

B. Updated E-Sample Identification Input File

Description: The input E-Sample Identification Input File is updated with the results of E-sample identification. HCUF HUs are assigned indicators denoting whether they are in or out of the E sample, whether or not they were initially eligible for the E sample, which weight should apply to each HU, to which segment each HU was assigned, and whether they are corresponding or not.

Level: Housing Unit

Scope: One record for each census HU in A.C.E. sample clusters.

Layout: See Attachment G.

VII. PROCESS

Apply the following steps to each block cluster on the E-Sample Identification Input File. Attachment B contains a flowchart of the E-sample identification process, and Attachment E gives a summary table of the process.

A. Determine the Number of HCUF HUs in the Block Cluster

1. Tally the number of HCUF HUs in the cluster, and denote this tally NHUCUF.
2. Obtain the American Indian Country Indicator (AICIND) for the cluster from the Sample Design File.
3. Determine which path the cluster should follow:
  - If  $NHUCUF \geq 80$  and  $AICIND \neq 1$ , then proceed to step B below.
  - If  $NHUCUF < 80$  or  $AICIND = 1$ , then assign the following variables to each HCUF HU in the cluster and proceed to step F where weights, summary counts, and other information will be determined.
    - a. E-Sample Indicator = 1
    - b. Preliminary E-sample Indicator = 0
    - c. E-Sample Probability Code = 1
    - d. Segment Identifier = AA
    - e. HU Group = 1
    - f. MSTATUS = 0

B. Determine if the Block Cluster is a Special Case Cluster

1. Obtain LEIND, RELIST, and NINT for the cluster from the Sample Design File.
2. Tally the following counts for the cluster and denote as indicated:
  - a. Number of SPEL HUs in the cluster, NHUEL
  - b. Number of SPEL HUs with after follow-up match codes of UI or CI, NUICI
3. If  $LEIND = 1$ ,  $RELIST = 1$ ,  $NINT = 0$ , or  $NHUEL = NUICI$ , then the block cluster is a special case cluster and continues to step B.4.  
  
Otherwise, the cluster is not a special case cluster. Proceed to step C below.
4. Tally the number of HCUF HUs in the cluster, and denote this tally NHUCUFS2.
5. Assign the following variables to each HCUF HU in the cluster:
  - a. Preliminary E-Sample Indicator = 1
  - b. E-Sample Probability Code = 2
  - c. Segment Identifier = AA
  - d. MSTATUS = 3
6. Proceed to step E below since these clusters require E-sample subsampling. Special case clusters with fewer than 80 HCUF HUs were screened out in step A.

C. Determine the E-Sample Status for HCUF HUs

1. Compare the CIDs of HCUF HUs to the CIDs assigned to SPEL HUs in the cluster. Corresponding HCUF HUs have their CIDs represented on the SPEL, while non-corresponding HCUF HUs do not.
  - For corresponding HCUF HUs, set MSTATUS = 1
  - For non-corresponding HCUF HUs, set MSTATUS = 2
2. Obtain INTERVW and DSSDSEG from the SPEL for each corresponding HU.

3. Do the following for each corresponding HCUF HU in the cluster:

- a. Assign E-Sample Probability Code = 1
- b. If INTERVW = 1 or 9, then assign the following variables:
  - i. E-Sample Indicator = 1
  - ii. Preliminary E-Sample Indicator = 0
  - iii. Segment Identifier = DSSDSEG
  - iv. HU Group = 2
- c. If INTERVW = 0 or 8, then assign the following variables:
  - i. E-Sample Indicator = 2
  - ii. Preliminary E-Sample Indicator = 0
  - iii. Segment Identifier = DSSDSEG
  - iv. HU Group = 3

4. For each non-corresponding HCUF HU in the cluster, assign E-Sample Probability Code = 2 and proceed to step D.

D. Determine if E-Sample Subsampling is Required in Non-Special Case Clusters

Tally the number of non-corresponding HCUF HUs in the block cluster. If this tally is zero, then E-sample subsampling is not required; proceed to step F below. Otherwise, do the following to determine if subsampling is required:

1. Obtain TEACR from the Sample Design File. Sort the non-corresponding HCUF HUs with the corresponding HCUF HUs in the cluster as follows:
  - For clusters with city-style addresses (TEACR = 1), sort by<sup>2</sup>:
    - Block and block suffix
    - Street name
    - House number
    - Unit designation
    - CID

---

<sup>2</sup>These variables are treated as character in the sort algorithm and thus the length of the character string factors into the sort. For example, APT 9 in unit designation sort will come before APT 1A because APT 1A has a longer character string.

- For clusters with non-city-style addresses (TEACR = 2), sort by:
  - Block and block suffix
  - Census map spot number
  - Within map spot number
  - CID
- 2. Assign to each non-corresponding HCUF HU in the cluster the segment identifier of the corresponding HCUF HU prior to it. If the first HCUF HU in the sort is non-corresponding, assign it the segment identifier of the last segment assigned to that block cluster during large block cluster subsampling. See Attachment F for an example of this assignment.
- 3. Set the E-sample indicators for non-corresponding HCUF HUs as follows:
  - If INTERVW = 1 or 9 for HUs in the same segment on the SPEL, then set Preliminary E-Sample Indicator = 1 for each non-corresponding HCUF HU in the cluster.
  - If INTERVW = 0 or 8 for HUs in the same segment on the SPEL, then assign the following variables for each non-corresponding HCUF HU in the cluster:
    - a. E-Sample Indicator = 2
    - b. Preliminary E-Sample Indicator = 0
    - c. HU Group = 7
- 4. Determine if E-sample subsampling is required in the cluster as follows:
  - Tally the number of HCUF HUs with a Preliminary E-Sample Indicator = 1 in the cluster and denote this tally NHUCUFS2.
  - If NHUCUFS2 < 80 then:
    - a. For HCUF HUs with Preliminary E-Sample Indicator = 1, assign E-Sample Indicator = 1
    - b. Assign HU Group = 4 for the same HCUF HUs
    - c. Proceed to step F below.
  - If NHUCUFS2 ≥ 80 then proceed to step E below.

E. E-sample Subsampling

Select a subsample of HCUF HUs with Preliminary E-Sample Indicator = 1 in a cluster as follows:

1. HUs in special case clusters have not yet been sorted. Obtain TEACR for the cluster from the Sample Design File. Sort the HCUF HUs with a Preliminary E-Sample Indicator = 1 as follows:
  - For clusters with city-style addresses (TEACR = 1), sort by<sup>3</sup>:
    - Block and block suffix
    - Street name
    - House number
    - Unit designation
    - CID
  - For clusters with non-city-style addresses (TEACR = 2), sort by:
    - Block and block suffix
    - Census map spot number
    - Within map spot number
    - CID
2. Calculate the take-every, TEES:
  - $$\text{TEES} = \frac{\text{NHUCUFS2}}{40}$$
  - If  $\text{TEES} > 4.000000$ , set  $\text{TEES} = 4.000000$ .
3. Assign each HCUF HU with Preliminary E-Sample Indicator = 1 an order number, ON, starting at one and incrementing by one until all such HUs in the cluster have an ON. The largest order number in the cluster will equal the value of NHUCUFS2.
4. Generate a sequence of numbers  $L_1, \dots, L_n$  as follows:
  - a. Generate a random number between 0 and 1 ( $0 < \text{RN} \leq 1$ ).

---

<sup>3</sup>These variables are treated as character in the sort algorithm and thus the length of the character string factors into the sort. For example, APT 9 in unit designation sort will come before APT 1A because APT 1A has a longer character string.

- b. Calculate a random start,  $RSES = RN \times TEES$ .
- c. Let  $L_1 = RSES$ .
- d. Calculate  $L_j = L_{j-1} + TEES$ , for  $j = 2, \dots, n$ , where  $n$  is the largest integer such that  $[RSES + (n - 1) \times TEES] \leq NHUCUFS2$ .
- e. Round each  $L_j$  up to the nearest integer (an integer rounds to itself).
- f. Each HCUF HU with Preliminary E-Sample Indicator = 1 and with ON equal to the rounded values of  $L_j$ ,  $j = 1, \dots, n$ , is in the E sample. Assign the following for each of these HUs:
  - i. E-Sample Indicator = 1
  - ii. For HUs in non-special case clusters, set HU Group = 5
  - iii. For HUs in special case clusters, set HU Group = 8
- g. Each HU with Preliminary E-Sample Indicator = 1 and with ON not equal to the rounded values of  $L_j$ ,  $j = 1, \dots, n$ , is not in the E sample. Assign the following for each of these HUs:
  - i. E-Sample Indicator = 2
  - ii. For HUs in non-special case clusters, set HU Group = 6
  - iii. For HUs in special case clusters, set HU Group = 9

For example, let  $NHUCUFS2 = 122$  and  $RN = 0.345167$ . Then  $TEES = 3.050000$ ,  $RSES = 1.052759$  and  $n = 40$ . Set  $L_1 = 1.052759$ . The generated  $L_j$ s would be the sequence: 1.052759, 4.102759, 7.152759, 10.202759, 13.252759, 16.302759, 19.352759, 22.402759, 25.452759, 28.502759, 31.552759, 34.602759, 37.652759, 40.702759, 43.752759, 46.802759, 49.852759, 52.902759, 55.952759, 59.002759, 62.052759, 65.102759, 68.152759, 71.202759, 74.252759, 77.302759, 80.352759, 83.402759, 86.452759, 89.502759, 92.552759, 95.602759, 98.652759, 101.702759, 104.752759, 107.802759, 110.852759, 113.902759, 116.952759 and 120.002759. Therefore, the HUs with ON values of 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32, 35, 38, 41, 44, 47, 50, 53, 56, 60, 63, 66, 69, 72, 75, 78, 81, 84, 87, 90, 93, 96, 99, 102, 105, 108, 111, 114, 117, and 121 would be selected for the sample.

5. Check the number of sampled HUs by calculating c:

$$c = \left| \frac{NHUCUFS2}{TEES} - n \right|$$



If the sampling is implemented correctly,  $c$  will be less than 1. For values of  $c$  that are greater than or equal to one, contact the Sample Design Team in the DSSD for review of the sampling operations.

F. Update Files

1. Sample Design File Version 6

a. HCUF and E-sample HU Counts

Compute nine HU counts for each cluster according to Table 1 below and record these totals to the Sample Design File. A value of 1 in the HU Group column in Table 1 indicates the counts in which an HCUF HU in that HU Group should be included. For example, an HCUF HU falling into HU group 1 is included in NHUES1, NHUES, NHUCUFS1, NHUCUFS, NHUCUF1, and NHUCUF. However, an HCUF HU falling into HU group 3 is only included in NHUCUF1 and NHUCUF.

Table 1. HCUF and E-Sample Housing Unit Counts

Count	HU Group								
	1	2	3	4	5	6	7	8	9
NHUES1	1	1	0	0	0	0	0	0	0
NHUES2	0	0	0	1	1	0	0	1	0
NHUES	1	1	0	1	1	0	0	1	0
NHUCUFS1	1	1	0	0	0	0	0	0	0
NHUCUFS2	0	0	0	1	1	1	0	1	1
NHUCUFS	1	1	0	1	1	1	0	1	1
NHUCUF1	1	1	1	0	0	0	0	0	0
NHUCUF2	0	0	0	1	1	1	1	1	1
NHUCUF	1	1	1	1	1	1	1	1	1

Notes:

1. In clusters with fewer than 80 HCUF HUs, NHUCUF1, NHUCUF, NHUCUFS1, NHUCUFS, NHUES1, and NHUES are all equal since all HCUF HUs are in the E sample. Also, NHUCUF2, NHUCUFS2, and NHUES2 are always equal to zero in these clusters since no HCUF HUs have an E-Sample Probability Code of 2.
2. In clusters with 80 or more HCUF HUs, NHUCUFS1 and NHUES1 are equal since all corresponding HCUF HUs in selected segments are in the E sample.

3. For special case clusters with 80 or more HCUF HUs, NHUCUF2 and NHUCUFS2 are equal since no segment boundaries can be defined. In addition, NHUCUF1 and NHUCUFS1 are always equal to zero in these clusters since there are no HCUF HUs with an E-Sample Probability Code of 1.

NHUES1, NHUES2, and NHUES are the number of HCUF HUs in the E sample with each E-Sample Probability Code.

- NHUES1: the number of E-sample HUs with E-Sample Probability Code = 1
- NHUES2: the number of E-sample HUs with E-Sample Probability Code = 2
- NHUES: the total number of E-sample HUs

NHUCUFS1, NHUCUFS2, and NHUCUFS are the number of HCUF HUs in segments selected for the P sample with each E-Sample Probability Code.

- NHUCUFS1: the number of HCUF HUs in selected segments with E-Sample Probability Code = 1
- NHUCUFS2: the number of HCUF HUs in selected segments with E-Sample Probability Code = 2
- NHUCUFS: the total number of HCUF HUs in selected segments

NHUCUF1, NHUCUF2, and NHUCUF are the number of HCUF HUs in the cluster with each E-Sample Probability Code.

- NHUCUF1: the number of HCUF HUs with E-Sample Probability Code = 1.
- NHUCUF2: the number of HCUF HUs with E-Sample Probability Code = 2.
- NHUCUF: the total number of HCUF HUs in the block cluster.

b. E-sample Weights

On the Sample Design File, there will be two E-sample cluster weights, WEIGHTE1 and WEIGHTE2. The HCUF HUs with an E-Sample Probability Code = 1 have a weight of WEIGHTE1, and the HCUF HUs with an E-Sample Probability Code = 2 get WEIGHTE2. If a cluster has only one type of HU Group, then only one weight will apply and the other weight will remain blank.

Record weights using the following criteria:

- i. If a cluster has HCUF HUs with HU Group = 1, set  $WEIGHTE1 = WEIGHTC$ , which is obtained from the Sample Design File Version 5.
- ii. If a cluster has HCUF HUs with HU Group = 2, set  $WEIGHTE1 = WEIGHTP$  from the Sample Design File Version 5.
- iii. If a cluster has HCUF HUs with HU Group = 4, set  $WEIGHTE2 = WEIGHTP$ .
- iv. If a cluster has HCUF HUs with HU Group = 5, set  $WEIGHTE2 = TE1 \times TE2 \times TEAR \times FTESB \times TELB \times TEES$ , where TE1, TE2, TEAR, FTESB, and TELB are obtained from the Sample Design File.
- v. If a cluster has HCUF HUs with HU Group = 8, set  $WEIGHTE2 = TE1 \times TE2 \times TEAR \times FTESB \times TEES$ .

c. E-Sample Identification Cluster Category

Set the E-Sample Identification Cluster Category (EICC) on the Sample Design File for each cluster using the following rules:

- If  $NHUCUF < 80$  then  $EICC = 1$
- If  $NHUCUF \geq 80$  and  $NHUCUFS < 80$  then  $EICC = 2$
- If  $NHUCUF \geq 80$  and  $NHUCUFS \geq 80$  then  $EICC = 3$
- If  $NHUCUF \geq 80$  and  $RELIST = 1$  then  $EICC = 4$
- If  $NHUCUF \geq 80$  and  $LEIND = 1$  then  $EICC = 5$
- If  $NHUCUF \geq 80$  and  $NHUEL > 0$  and  $NHUEL = NUICI$  then  $EICC = 6$
- If  $NHUCUF \geq 80$  and  $NHUEL = 0$  then  $EICC = 7$

d. E-Sample Subsampling Information

Record the following variables for each cluster. If E-sample subsampling was not required in a cluster, set both variables equal to one.

- E-sample subsampling random start, RSES
- E-sample subsampling take-every, TEES

## 2. E-Sample Identification Input File

Update the E-Sample Identification Input File with the results of E-sample identification. The layout of this file is in Attachment G. Specifically, record the following variables for each HCUF HU:

- E-Sample Indicator
- E-Sample Probability Code
- Segment Identifier
- MSTATUS, the HCUF HU correspondence indicator

## VIII. VERIFICATION

Verification of the E-sample identification will include both micro-level independent replication of the process and macro-level analysis of to-be-determined summary statistics. The Sample Design Team in the DSSD will perform all activities involved in both types of verification. Access to the following files is required:

- Sample Design File Version 5
- Sample Design File Version 6
- Subsampled Preliminary Enhanced List
- Updated E-Sample Identification Input File

## IX. REFERENCES

- 1 DSSD Census 2000 Procedures and Operations Memorandum Series R-31, "Accuracy and Coverage Evaluation Survey: Specification for E-Sample Identification," May 16, 2000.
- 2 DSSD Census 2000 Procedures and Operations Memorandum Series R-37, "Accuracy and Coverage Evaluation Survey: E-Sample Identification Approval and Summary of Results," November 17, 2000.
- 3 DSSD Census 2000 Procedures and Operations Memorandum Series Chapter S-HU-08, "Creation of the Census 2000 Accuracy and Coverage Evaluation (A.C.E.) Enhanced List for Person Phase Interviewing," December 5, 2000.
- 4 DSSD Census 2000 Procedures and Operations Memorandum Series R-27, "Accuracy and Coverage Evaluation: Large Block Cluster Subsampling Specifications," March 8, 2000.

5 DSSD Census 2000 Procedures and Operations Memorandum Series Chapter  
S-DT-01, "Accuracy and Coverage Evaluation: The Design Document,"  
January 11, 2000.

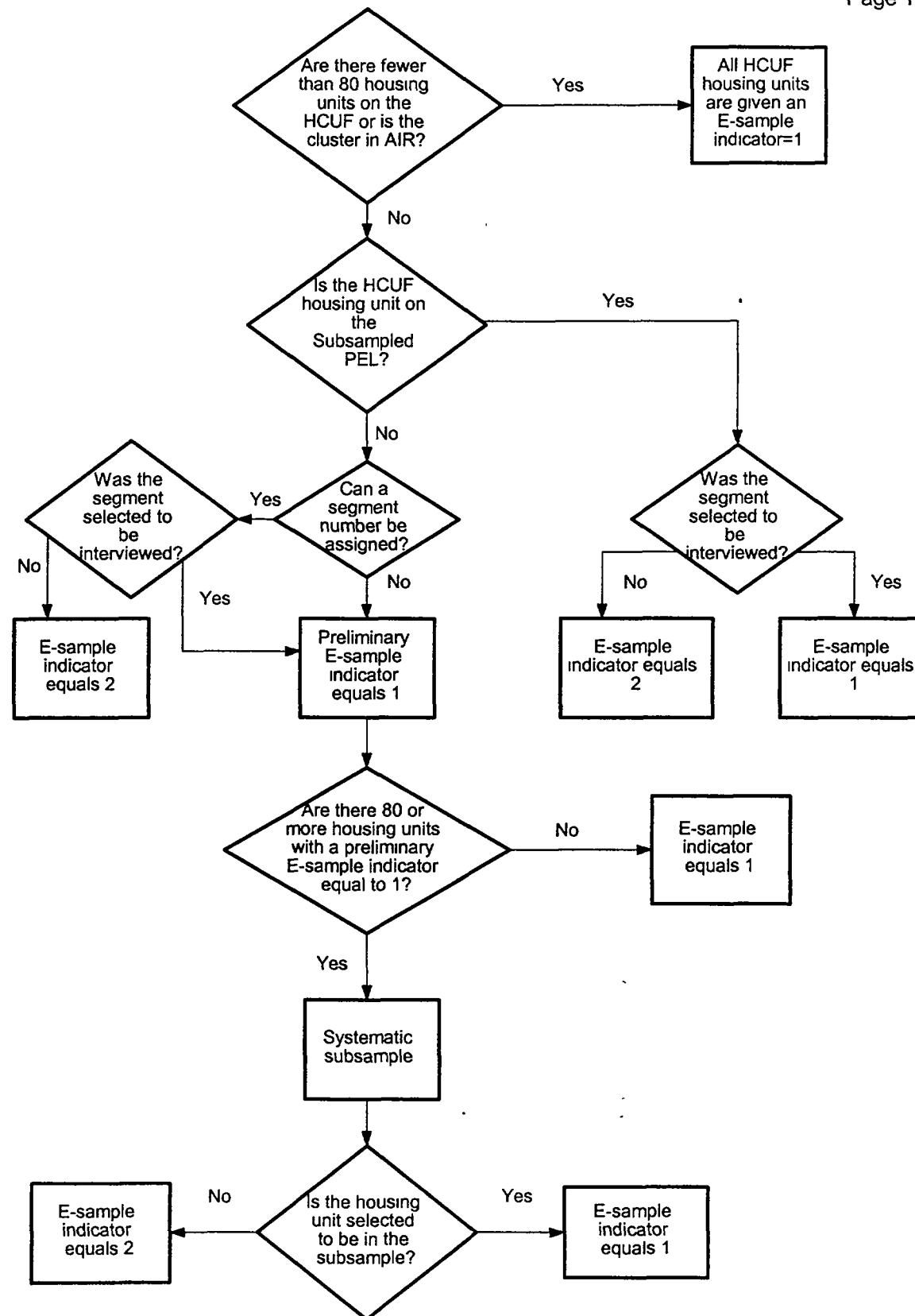
cc. DSSD Census 2000 Procedures and Operations Memorandum Series Distribution List  
A.C.E. Implementation Team  
Statistical Design Team Leaders  
Sample Design Team

Description of HU Groups

HU Group Code	Description
1	HCUF HUs in clusters with < 80 HCUF HUs or in AIR clusters
2	Corresponding HCUF HUs in clusters with 80+ HCUF HUs and in segments selected for the P sample
3	Corresponding HCUF HUs in 80+ clusters and in segments not selected for the P sample
4	Non-corresponding HCUF HUs in 80+ clusters, in P-sample segments, and with no E-sample subsampling
5	Non-corresponding HCUF HUs in clusters with 80+ HCUF HUs, in P-sample segments, and selected in E-sample subsampling
6	Non-corresponding HCUF HUs in 80+ clusters, in P-sample segments, and not selected in E-sample subsampling
7	Non-corresponding HCUF HUs in 80+ clusters and in segments not selected for the P sample
8	HCUF HUs in 80+ special case clusters and selected in E-sample subsampling
9	HCUF HUs in 80+ special case clusters and not selected in E-sample subsampling

# Flowchart of the E-sample Identification Process

Attachment B  
Page 1 of 1



## Layout of the Subsampled Preliminary Enhanced List

Layout Name : ENHANCED00.LAY  
 Description : 2000 ENHANCED LIST LAYOUT  
 Total Length : 360  
 Date Created : 05-01-2000

Page : 1

#	Field	Field description	length	Positions Beg - End	
1.	CNTRLNM	CONTROL NUMBER	24	1 - 24	CHAR
		1: 4 LCO			
		5:10 CLUSTER			
		11:12 SEGMENT			
		13:17 MAP SPOT NUMBER			
		18:21 WITHIN MSN ID			
		22:24 ZERO FILL			
2.	LCO	LOCAL CENSUS OFFICE	4	25 - 28	CHAR
		*****			
		Index 1 CLUST thru WMSN			
		*****			
3.	CLUST	CLUSTER NUMBER	6	29 - 34	CHAR
4.	MSN	ENHANCED IL MAP SPOT NUMBER	5	35 - 39	CHAR
5.	WMSN	WITHIN MAP SPOT NUMBER ID	4	40 - 43	CHAR
		*****			
		Index 2 CID			
		*****			
6.	CID	MAF ID	12	44 - 55	CHAR
7.	BLK	1998 BLOCK AND SUFFIX	6	56 - 61	CHAR
8.	URBNZ	URBANIZATION	30	62 - 91	CHAR
9.	HSNUM	HOUSE NUMBER (LJ/BF)	10	92 - 101	CHAR
10.	SNAME	STREET NAME (LJ/BF)	35	102 - 136	CHAR
11.	UNIT	UNIT DESIGNATION (LJ/BF)	15	137 - 151	CHAR
12.	RR	RURAL ROUTE/BOX # (LJ/BF)	25	152 - 176	CHAR
13.	POBX	PO BOX NUMBER (LJ/BF)	10	177 - 186	CHAR
14.	CITY	CITY/TOWN NAME	20	187 - 206	CHAR
15.	ZIP	ZIP CODE	5	207 - 211	CHAR
16.	ZIP4	ZIP + 4	4	212 - 215	CHAR
17.	STATE	FIPS STATE ABBREVIATION	2	216 - 217	CHAR
18.	FIPSCNTY	FIPS COUNTY CODE	3	218 - 220	CHAR
19.	FIPST	FIPS STATE CODE	2	221 - 222	CHAR
20.	PL	PHYSICAL LOCATION DESCRIPTION	50	223 - 272	CHAR
21.	PRKNM	TRAILER PARK NAME	30	273 - 302	CHAR
22.	HUFIN	MATCH CODE FROM HU MATCHING	2	303 - 304	CHAR
23.	HUFINID	ID FROM HOUSING UNIT MATCHING	12	305 - 316	CHAR
24.	TOA	TYPE OF BASIC ADDRESS	1	317 - 317	CHAR
		1 = ONE FAMILY HOUSE			
		2 = BSA WITH 2 OR MORE HUS			
		3 = MOBILE HOME NOT IN PARK			
		4 = MOBILE HOME IN PARK			
		5 = ONE FAMILY HOME IN			
		SPECIAL PLACE			
		6 = BSA WITH 2 OR MORE HUS			
		IN A SPECIAL PLACE			
		7 = OTHER			
25.	USTAT	UNIT STATUS	1	318 - 318	CHAR
		1 = OCCUPIED OR VACANT AND			
		INTENDED FOR OCCUPANCY			
		2 = UNDER CONSTRUCTION			
		3 = FUTURE CONSTRUCTION			
		4 = UNFIT FOR HABITATION			
		5 = BOARDED UP			
		6 = STORAGE OF HOUSEHOLD			



Layout Name : ENHANCED00.LAY  
Description : 2000 ENHANCED LIST LAYOUT  
Total Length : 360  
Date Created : 05-01-2000

Page : 2

#	Field	Field description	length	Positions	
				Beg	End
		GOODS			
		7 = VACANT MOBILE HOME SITE			
		8 = OTHER			
26.	UR	(Not used in 2000)	1	319	319 CHAR
27.	QAFLG	QA SAMPLE FLAG	1	320	320 CHAR
		0 = NOT IN QA SAMPLE			
		1 = IN QA SAMPLE			
28.	ESAMPFLG	E-SAMPLE ELIGIBILITY FLAG	1	321	321 CHAR
29.	URFLAG	FLAG INDICATING THAT ADDRESS IS CONSIDERED TO BE URBAN OR RURAL	1	322	322 CHAR
		0 = RURAL			
		1 = URBAN			
30.	MULTIFLAG	FLAG INDICATING THAT UNIT IS IN A MULTIUNIT OF LESS THAN 20 UNITS	1	323	323 CHAR
		0 = MULTI <20 UNITS			
		1 = NONMULTI, OR MULTI >= 20			
31.	DSSDSEG	SEGMENT FOR LARGE BLOCK SUBSAM	2	324	325 CHAR
32.	FLDSEG	SEGMENT FOR ASSIGNING WORK IN	2	326	327 CHAR
33.	INTERVW	AFTER LARGE BLOCK SUBSAMP	1	328	328 CHAR
		0 = OUT OF SAMPLE			
		1 = IN SAMPLE			
		8 = SUPP/OUT OF SAMPLE			
		9 = SUPP/IN SAMPLE			
34.	JIC	JUST IN CASE SPACE ***** THESE FIELDS ARE USED FOR LARG BLOCK SUBSAMPLING. *****	14	329	342 CHAR
35.	TOTCASES	NUMBER OF CASES IN CLUSTER	6	343	348 CHAR
36.	ICMCASES	NUMBER OF ICM CASES IN CLUSTER	6	349	354 CHAR
37.	CENCASES	NUMBER OF CEN CASES IN CLUSTER	6	355	360 CHAR

## Layout of the Sample Design File

<u>Variable Description</u>	<u>Name</u>	<u>Places</u>	<u>Source</u>
Census Region	REGION	1	UN
Census Division	DIV	2	UN
State code	STATE	3-4	UN
County code	COUNTY	5-7	UN
Local census office	LCO	8-11	CS
Interim Tract (Pseudo Tract)	ITRACT	12-17	BC
Current Sample Indicator	CSI	19	UO
A.C.E. block cluster number	CLUST	21-25	CS
Check Digit	DIGIT	26	CS
Geography block cluster number	GCLUST	28-32	BC
List/Enumerate Indicator	LEIND	33	BC
Type of Enumeration Area Recode	TEACR	34	CS
Type of Enumeration Area Group Revised	TEAGR	35	RV
Type of Enumeration Area group	TEAG	36	BC
Number of HUs used for sample design	NHU	37-41	BC
Number of MAF HUs	NHUM	43-47	BC
Number of 1990 HUs	NHU90	49-53	BC
Sampling Stratum	SS	55	UN
1 = Small			
2 = Medium			
3 = Large			
4 = American Indian Reservation			
American Indian Country Indicator	AICIND	56	BC
0 = No American Indian Country			
1 = American Indian Reservation/trust land			
2 = Tribal Jurisdiction Area/ Alaska Native Village Statistical Area/ Tribal Designated Statistical Area			
Demographic/Tenure Group code	DTCODE	57-58	UN
Demographic/Tenure Group label	DTLABEL	59-60	UN
Estimated Urbanicity of block cluster	ECLUSURB	62	UN
1 = Urban Area with population $\geq 250,000$			
2 = Other Urban Area			
3 = Non-Urban Area			
Size Category	SIZCAT	63	UN
1=Small (0-2 hus)			
2=Medium (3-79 hus)			
3=Large (80+ hus)			
Additional space		64-91	

<u>Variable Description</u>	<u>Name</u>	<u>Places</u>	<u>Source</u>
First step index number	INDEX1	92-99	CS
Listing sample selection indicator 1 = Selected	BC1	101	CS
Random Start for listing sample selection	RS1	103-113	UN
Take-every for listing sample selection	TE1	115-125	UN
Second step listing sample selection indicator 0 = Not Selected 1 = Selected	BC2	127	CS
Random Start for the second step of the listing sampling	RS2	129-139	CS
Take-every for the second step of the listing sampling	TE2	141-151	CS
Unbiased weight after block cluster sampling	WEIGHTBC	153-164	CS
Additional space		165-175	
Preliminary Number of HUs on the Independent List	NHUILP	176-180	AR
Number of Housing Units On the January 2000 DMAF	NHUDMAF	182-186	AR
Demographic Code 1 = Minority 2 = Non-Minority 3 = Puerto-Rico	DEMCODE	188	AR
Consistency Code 1 = Low Inconsistent (IL significantly smaller than DMAF) 2 = Consistent 3 = High Inconsistent ((IL significantly larger than DMAF)	CONCODE	189	AR
A.C.E. Reduction Stratum	ARST	190-191	AR
A.C.E. Reduction Indicator 0 = Not Selected 1 = Selected	ACERED	193	AR
Random Start for A.C.E. Reduction	RSAR	195-205	AR
Take-every for A.C.E. Reduction	TEAR	207-217	AR
Unbiased weight after A.C.E. reduction	WEIGHTAR	219-230	AR
Collapsing Flag	COLFLAG	232	AR
A.C.E. Reduction Index Number	INDEXR	234-241	AR
Number of Housing Units On the December 1999 DMAF (Initial)	NHUDMAFI	243-247	AR
Additional space		248-294	
Number of HUs on the Independent List Revised (equals NHUIL for all but 4 clusters (see memo R-34))	NHUILR	295-299	RV
Number of HUs on the Independent List	NHUIL	301-305	SB
Small Block Cluster Subsampling Stratum	SBCSS	306-307	SB
Small Block Subsampling Indicator 0 = Not Selected 1 = Selected	SB	308	SB
Random Start for Small Block subsampling	RSSB	310-320	SB
Initial take-every for Small Block subsampling	ITESB	322-332	SB
Unbiased weight for A.C.E. cluster	WEIGHTC	334-345	SB
Larger of the DMAF and IL HU count	LARGERHU	347-351	SB
Final take-every for Small Block subsampling	FTESB	352-362	SB
Additional space		363-370	

<u>Variable Description</u>	<u>Name</u>	<u>Places</u>	<u>Source</u>
Relisted Block Cluster Flag 0 = Not Relisted, 1 = Relisted	RELIST	371	LB
Number of total hus in block cluster	NHUEL	373-377	LB
Number of A.C.E. hus in cluster	NHUELA	379-383	LB
Number of supplemental hus in cluster	NHUELN	385-389	LB
Large Block Cluster EL subsampling code 1 = NHUELA < 80 hus, 2 = NHUELA ≥ 80 hus	ELLBSUB	391	LB
Random Start for Large Block subsampling	RSLB	393-403	LB
Take-every for Large Block subsampling	TELB	405-415	LB
Number of segments in block cluster	NSEG	417-418	LB
Number of segments selected in block cluster	NSEGSAM	420-421	LB
Day of Arrival	DAY	423-424	LB
Final Cluster Order Number	CON	431-434	LB
Number of total hus for interview in block cluster	NINT	436-440	LB
Unbiased weight for P-sample HUs	WEIGHTP	442-453	LB
Number of Assignments in block cluster	NA	455-456	LB
Final Sampling Strata	FSS	458-464	LB
Additional space		465-490	
Number of HCUF HUs in block cluster with an ESPS code of 1	NHUCUF1	491-495	ES
Number of HCUF HUs in block cluster with an ESPS code of 2	NHUCUF2	497-501	ES
Number of HCUF HUs in block cluster	NHUCUF	503-507	ES
Number of HCUF HUs in selected segments with an ESPS code of 1	NHUCUFS1	509-513	ES
Number of HCUF HUs in selected segments with an ESPS code of 2	NHUCUFS2	515-519	ES
Number of HCUF HUs in selected segments of a block cluster	NHUCUFS	521-525	ES
E-Sample Identification cluster category 1 = NHUCUF < 80 2 = NHUCUF ≥ 80 and NHUCUFS < 80 3 = NHUCUF ≥ 80 and NHUCUFS ≥ 80 4 = NHUCUF ≥ 80 and RELIST = 1 5 = NHUCUF ≥ 80 and List/Enumerate 6 = NHUCUF ≥ 80 and only UI/CI PEL HUs 7 = NHUCUF ≥ 80 and zero PEL HUs	EICC	527	ES
Random Start for E-sample subsampling	RSES	529-539	ES
Take-every for E-sample subsampling	TEES	541-551	ES
Number of E-sample HUs in block cluster with an ESPS code of 1	NHUES1	553-557	ES
Number of E-sample HUs in block cluster with an ESPS code of 2	NHUES2	559-563	ES
Number of E-sample HUs in block cluster	NHUES	565-569	ES
Unbiased weight for E-sample HUs with an ESPS code of 1	WEIGHTE1	571-582	ES
Unbiased weight for E-sample HUs with an ESPS code of 2	WEIGHTE2	584-595	ES

<u>Variable Description</u>	<u>Name</u>	<u>Places</u>	<u>Source</u>
Number of confirmed A.C.E. housing units not found in the census	CURCI	676-680	TES
Number of unconfirmed A.C.E. housing units not found in the census	CURUI	682-686	TES
Number of census housing units geocoded to the wrong census block	CURGE	688-692	TES
Targeted extended search selection type	TESSELECT	694	TES
Targeted extended search selection flag	TESFLAG	696	TES
Random Start for the targeted extended search	RSTES	698-709	TES
Take-every for the targeted extended search	TETES	710-721	TES
Targeted Extended Search Index Number	TESN	722-727	TES
Additional Space		728-750	

Source Codes

AR: A.C.E. Reduction  
 BC: Block Clustering  
 CS: Block Cluster Sampling  
 ES: E-sample Identification  
 LB: Large Block Subsampling  
 SB: Small Block Subsampling  
 UN: Universe File Creation  
 UO: Updated for each operation  
 TES: Targeted Extended Search  
 RV: Variable Revision

See the following page for notes on this table.

Notes:

1. In HU Groups 5 and 8, “Samp” means the HCUF HU was selected during E-sample subsampling. Likewise “Not Samp” in HU Groups 6 and 9 means the HCUF HU was not selected.
2. WEIGHTE2 in HU Group 5 is  $TE1 \times TE2 \times TEAR \times FTESB \times TELB \times TEES$ . WEIGHTE2 in HU Group 8 is  $TE1 \times TE2 \times TEAR \times FTESB \times TEES$ .
3. WEIGHTC is the A.C.E. block cluster weight after small block cluster subsampling. WEIGHTP is the A.C.E. block cluster weight after large block cluster subsampling. TEES is the take-every used for E-sample subsampling. TE1, TE2, TEAR, FTESB, and TELB are take-every's from previous sampling operations and are obtained from the Sample Design File.
4. Possible combinations of weights within a block cluster are
  - All HUs have WEIGHTC
  - All HUs have WEIGHTP
  - All HUs have  $WEIGHTP \times TEES$
  - All HUs have  $WEIGHTC \times TEES$
  - Corresponding HUs have  $WEIGHTE1 = WEIGHTP$  and non-corresponding HUs have  $WEIGHTE2 = TE1 \times TE2 \times TEAR \times FTESB \times TELB \times TEES$
  - Corresponding HUs have  $WEIGHTE1 = WEIGHTP$  and non-corresponding HUs have  $WEIGHTE2 = TE1 \times TE2 \times TEAR \times FTESB \times TEES$
5. If E-Sample Probability Code = 1 use the WEIGHTE1 variable on the Sample Design File to apply the appropriate weight  
 = 2 use the WEIGHTE2 variable on the Sample Design File to apply the appropriate weight.
6. In HU Groups 3, 6, 7, and 9, “NA” means an E-sample weight is not applicable because the HU Group is not in the E sample.

### Example of Segment Identifier Assignment

The table on the next page contains an example of a block cluster with 88 HCUF HUs that correspond to HUs on the SPEL. This is a large cluster, but only a few of the HUs are shown for illustrative purposes.

The HCUF HUs that correspond with HUs on the SPEL have a segment identifier. The non-corresponding HCUF HUs are sorted into their proper order using the house number and street name. They are assigned the segment identifier of the previous corresponding HCUF HU. The segment(s) selected for the A.C.E. interview sample is also the segment(s) selected for the E sample to achieve an overlapping E sample and P sample. The corresponding HCUF HUs in selected segments are assigned an E-sample indicator equal to one. The non-corresponding HCUF HUs in selected segments are assigned a preliminary E-sample indicator equal to one and may be subject to subsampling if there are 80 or more of them. The HCUF HUs in non-selected segments are assigned an E-sample indicator equal to two.



Segment Identifier Example in a Block Cluster with City-Style Addresses

SPEL A.C.E. Map Spot Number	Segment Identifier from SPEL	Segment Identifier Assigned during E-sample Identification	Address
11	BA	CA	101 1st St
12	BA		102 1st St
13	CA		103 1st St
			105 1st St
4	AA	AA	309 Maple Ln
5	AA		311 Maple Ln
			311 Maple Ln Basement
7	BA		315 Maple Ln
1	AA	AA	104 Peach Ct
			110 Peach Ct
			112 Peach Ct
			116 Peach Ct
		AA	120 Peach Ct
2	AA		702 Vermont Ave
			704 Vermont Ave
3	AA		704 Vermont Ave rear
		AA	704 Vermont Ave Upper
.	.	.	.
.	.	.	.
.	.	.	.

Layout Name : ACE 2000 HCUF EXTRACT FOR E-SAMPLE IDENTIFICATION Page : 1  
 Description : ACELAY:HCUF\_ESAMPID.LAY  
 Total Length : 164  
 Date Created : 05-10-2000

#	Field	Field description	length	Positions		
				Beg	-	End
1.	LCO	LOCAL CENSUS OFFICE	4	1	-	4 CHAR
2.	ST	COLLECTION FIPS STATE CODE	2	5	-	6 CHAR
3.	COU	COLLECTION FIPS COUNTY CODE	3	7	-	9 CHAR
4.	BLK	BLOCK NUMBER	6	10	-	15 CHAR
		BLK(1:5) NUMERIC RJ,BF				
		BLK(6:6) ALPHA BF				
		* INDEX: CLUST/CID *				
5.	CLUST	CLUSTER NUMBER	6	16	-	21 CHAR
6.	CID	CENSUS UNIT ID	12	22	-	33 CHAR
7.	MSN	CENSUS MSN	5	34	-	38 CHAR
8.	UNIT	UNIT DESIGNATION RJ,BF	13	39	-	51 CHAR
9.	HSNUM	HOUSE NUMBER <HW>LJ	21	52	-	72 CHAR
10.	SNAME	STREET NAME <SW> LJ,BF	43	73	-	115 CHAR
11.	RR	RURAL ROUTE	8	116	-	123 CHAR
12.	BOX	BOX NUMBER/PO BOX	12	124	-	135 CHAR
13.	ZIP	ZIP CODE	5	136	-	140 CHAR
14.	ZIP4	ZIP + 4	4	141	-	144 CHAR
15.	ESAMP	E-SAMPLE INDICATOR	1	145	-	145 CHAR
		1 = ESAMPLE				
		2 = NON ESAMPLE				
16.	EPROB	E-SAMPLE PROBABILITY	1	146	-	146 CHAR
17.	MSTATUS	MATCH STATUS	1	147	-	147 CHAR
		1 = MATCH				
		2 = NONMATCH				
		3 = NOT APPLICABLE (UI/CI -				
		2 LARGE, RELIST, & A FEW				
		OTHER 100% UI/CI CLUSTERS				
18.	SEGMENT	SEGMENT FROM LB SUBSAMPLING	2	148	-	149 CHAR
19.	OUTFILL	FILLER FOR RETURN DATA	15	150	-	164 CHAR